Journal of Radioanalytical and Nuclear Chemistry, Vol. 249, No. 1 (2001) 121–131

Intercomparison study of inductively coupled plasma mass spectrometry, thermal ionization mass spectrometry and fission track analysis of μBq quantities of ^{239}Pu in synthetic urine

K. G. W. Inn,[1] D. McCurdy,[2] L. Kuruvilla,[2] N. M. Barss,[3] R. Pietrzak,[4] E. Kaplan,[4] W. Inkret,[5] W. Efurd,[5] D. Rokop,[5] D. Lewis,[5] P. Gautier,[5] R T. Bell III[3]

[1] National Institute of Standards and Technology, Gaithersburg, MD. USA
'Duke Engineering and Services. 580 Main Street, Bolton, MA. USA
[3] U.S. Department of Energy, Office of International Health Programs, Germantown, MD, USA
[4] Brookhaven National Laboratory, Building 703M, P.O. Box 5000, Upton, NY, USA
[5] Los Alamos National Laboratory Los Alamos, NM, USA

(Received December 13.2000)

Even today, some Marshall Islanders are looking forward to permanently resettling their islands after five decades. The U.S. Department of Energy and the resettled residents require reasonable but cost-prudent assurance that tbe doses to residents from residual ^{239}Pu will not exceed recognized international standards or recommendations, as estimated from the excretion of ^{239}Pu in urine. The goal of this study was to evaluate the bias, uncertainty and sensitivity of analytical techniques that measure 3–56 μBq ^{239}Pu in synthetic urine. The analytical techniques studied in this work included inductively coupled plasma mass spectrometry, thermal ionization mass spectrometry and fission track analysis. The results of the intercomparison demonstrated that all three techniques were capable of making the measurements, although not with equal degree of bias and uncertainty. The estimated minimum detectable activity was 1 μBq of ^{239}Pu per synthetic urine sample. This exercise is also the first effort to certify test materials of plutonium in the $nBq \cdot g^{-1}$ range.

Introduction

The Department of Energy (DOE), Office of International Health Programs (EH-63), is in the process of assisting Marshall Islanders to permanently resettle their islands after five decades. The DOE and the resettled residents require reasonable and cost-prudent assurances that the radiation dose to residents due to residual radioactivity from previous atmospheric nuclear weapons testing will not exceed recognized international standards or recommendations. One of the remaining radionuclides that could conuibute to the internal radiation dose from the inhalation and ingestion intake pathways is ^{239}Pu. Since biological samples can be collected to quantitate the body content of radioactive materials, the measurement of these parameters by instruments or analytical techniques must be accurately known. The internal uptake and the amount of ^{239}Pu in the body can be estimated from the excretion of ^{239}Pu in the urine of an individual. The analytical technique must typically have sufficient sensitivity to quantify ^{239}Pu at a level equivalent to 1 mSv per year with known statistical confidence that considers internal dose assessment uncertainties.

At the request of community leaders to monitor ^{239}Pu in the urine of individuals resettled at Enewetak Atoll, DOE has used fission track analysis (FTA) since the late 1980's. Although PTA is very sensitive, it is also expensive and requires long sample processing turnaround times. Thermal ionization mass spectrometry (TIMS) has had the potential sensitivity to equal that of FTA and could also provide isotopic information, but has never seriously been used for routine radiobioassay work because of the necessary labor-intensive and time-consuming chemical purifications and rare expertise to knowledgeably operate the instrument. More recent technology, inductively coupled plasma mass spectrometry (ICP-MS), offers great potential as a rapid alternative ultra-sensitive measurement method that is easy to operate and only requires a minimum amount of sample preparation. The attraction of faster, less expensive analyses of very low levels of plutonium in urine a t comparable sensitivity motivated the Department of Energy to assess the capabilities of all three of these measurement techniques.

The DOE selected two of its national laboratories to participate in this measurement capability demonstration study. The goals of this study were to evaluate the measurement techniques' bias, uncertainty and sensitivity mass in the concentration range of 3-56 μBq for 200 g samples of synthetic urine by: (1) comparing the Brookhaven National Laboratory's (BNL) inductively coupled plasma mass spectrometry (ICP-MS) with (2) BNL's fission track analysis (PTA); and (3) routine bioassay analysis methodology at the Los Alamos National Laboratory (LANL) that uses thermal ionization spectrometry (TIMS). The concentration range was specifically chosen to evaluate the methods' bias and uncertainty, and limits of detection. A priori estimates of measurement capabilities indicated that confidence varied among the measurement methodologies, depending on the particular objective that they were addressing. BNL had optimized the FTA method for low-level analysis of ^{239}Pu in mine and

0236–5731/2001/USD 17.00

Akadémiai Kiadó. Budapest
Kluwer Academic Publishers. Dordrecht

believed that their best results would be for the 18.5 and 46.3 nBq/g concentrations. Since LANL TIMS was focusing on routine response methodologies, their best results were predicted to be for the 148 and 278 nBq/g concentration samples. On the other hand, BNL had optimized their ICP-MS methodology for the 18.5 to 278 nBq/g concentrations and believed that their results would be consistent over the whole concentration range.

The major portion of the preparation tasks was performed by the Duke Engineering and Services Laboratory (DESL),* in terms of establishing the stability of ^{99m}Tc tracer for ^{239}Pu in the synthetic urine, executing the dilutions, conducting confirmational measurements, documenting the preparation work and distributing the samples to participating laboratories. NIST oversaw the development of the work plan and DESL's preparation of the test materials, certified the test materials and evaluated the resulting data.

Experimental

Approach

The dilution and measurement confirmation scheme (Fig. 1) was developed for the production of five replicate ^{239}Pu in synthetic urine at blank, 18.5, 46.3, **148.4,** and 277.7 $nBq \cdot g^{-1}$ samples for the three participating laboratories.

The ^{239}Pu test solutions were prepared from NIST Standard Reference Material (SRM) 4330A. SRM 4330A contained 37.9 $Bq \cdot g^{-1}$ ^{239}Pu, with negligible quantities of any isotopic or radiometric interferences.

The ^{239}Pu spiked synthetic urine samples (see Table 1 for component composition) were prepared, the dilutions were confirmed by isotope dilution alpha spectrometry and ^{99m}Tc tracer gamma-spectrometry, and five replicate samples at each concentration were distributed blind to BNL and LANL.

The participating laboratories had two months to report their final measurement data (including negative values) to NIST along with their evaluation of the propagated uncertainties in their measurements.

The resulting data were evaluated to determine the bias, uncertainty, sensitivity and limitations of the analyses of ^{239}Pu in synthetic urine for the following:

Individual ***laboratory results:*** *(normality tests):* data distribution (test for measurement control), mean value (bias), variance (uncertainty), potential measurement discrepancies.

Measurement discrepancies and resolution: measurement methodologies, sources of discrepancies and outlying data, outlying data evaluation.

Intercomparison performances: data distribution (normality tests), mean value (bias), variance (uncertainty).

Technology evaluation: bias, uncertainty, minimum detection amount.

Technical issues

A number of technical issues were raised during the design of the intercompatison protocols. These included: stability of the plutonium in glass bottles; stability of the plutonium in the synthetic urine; contamination from plutonium in the reagents used to make the synthetic urine; and adequacy of the synthetic urine as a substitute for natural urine.

Previous experience within the in vitro radiobioassay community indicated no particular problems with these issues. However, the issues were reassessed in this intercomparison because of the extremely low concentrations of plutonium. These issues were evaluated using the intercompatison data.

Table 1. Recipe for synthetic mine

Component	g/kg
Urea	***16.00***
NaCl	***2.32***
KC1	***3.43***
Creatinine	1.10
Na_2SO_4 (anhydrous)	***4.3*** 1
Hippuric acid	***0.63***
NH_4Cl	1.06
Citric acid	***0.54***
$MgSO_4$ (anhydrous)	***0.46***
$NaH_2PO_4 \cdot H_2O$	***2.73***
$CaCl_2 \cdot 2H_2O$	***0.63***
Oxalic acid	***0.02***
Lactic acid	***0.094***
Glucose	***0.48***
$Na_2SiO_3 \cdot 9H_2O$	0.071
Pepsin	***0.029***
Conc. nitric acid	***50.00***
'Yellow food color (optional)	***0.06***

* Formerly Yankee Atomic Environmental Laboratory, Westboro, MA, USA.

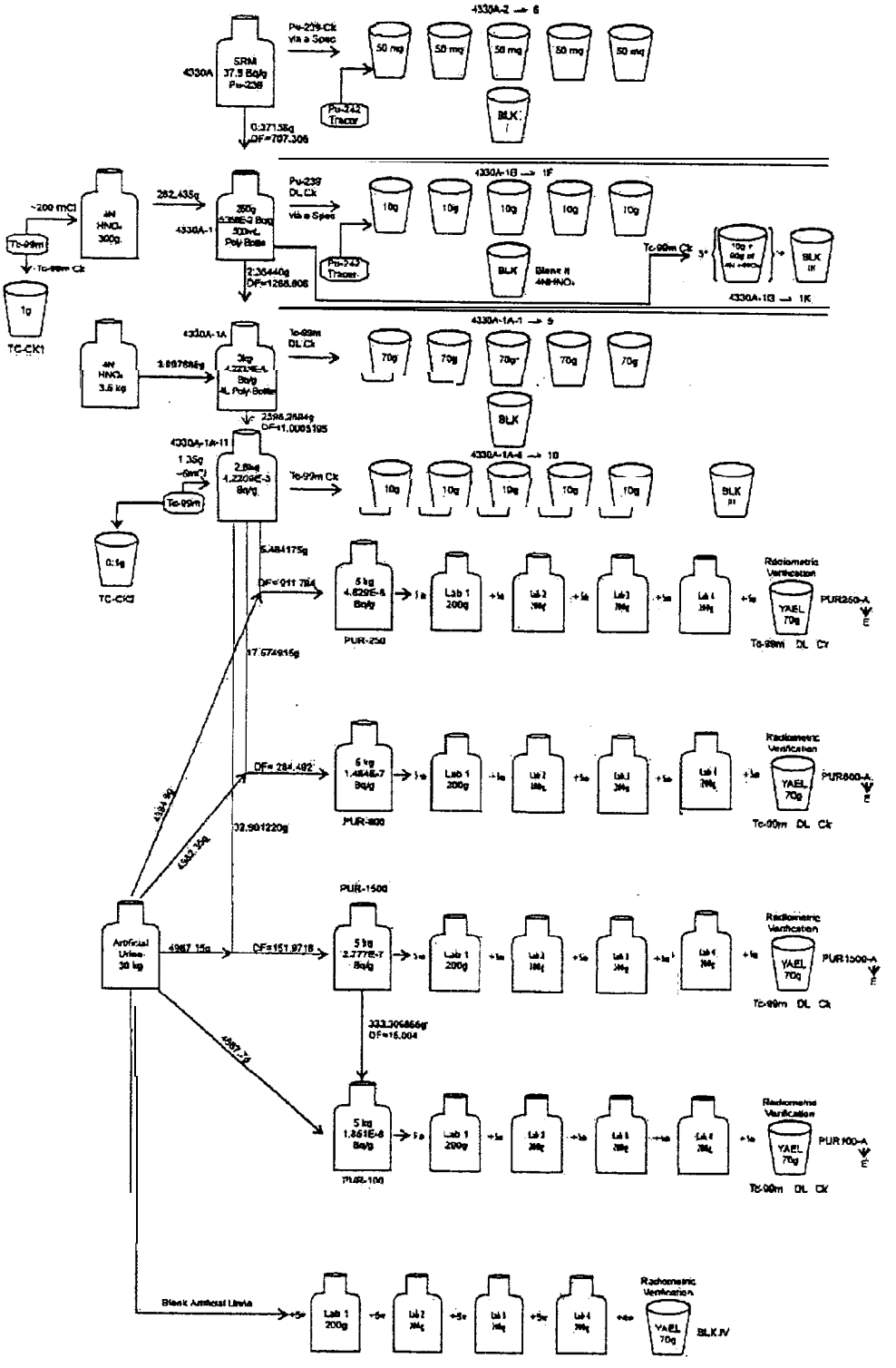

Fig. 1. Dilution and verification scheme for the production of the intercomparison materials

Table 2. Verification of dilutions

Dilution	Number of replicate sources	Difference between measured and gravimetric dilution factors, %	Relative combined uncertainty (lo), %	$t_{0.975}$ test, %
SRM 4330A to 4330A-1	5	-3.11	2.62	±3.26
4330A-1 to 4330A-1A	5	-2.58	3.21	k3.98
4330A-1A-11 to PUR1500	5	-2.81	2.57	±3.19
PUR1500 to PUR100	5	-1.87	2.58	±3.21
4330A-1A-11 to PUR250	5	-3.76	2.60	±3.23
4330A-1A-11 to PUR800	5	-3.03	2.50	Zt3.11

Table 3. Relative expanded uncertainty for each dilution

Sample identification number	Uncertainty components: Reference material uncertainty, (Is) %	Gravimetric, (1s) %	Relative combined uncertainty, (1s) %	Relative expanded uncertainty, (2s) %
SRM 4330A	0.36		0.36	0.7
4330A-1	0.36	0.14	0.4	0.8
4330A-1A	0.39	0.14	0.4	0.8
4330A-1A-11	0.41	0.14	0.4	0.9
PUR1500	0.44	0.14	0.5	1.0
PUR100	0.46	0.14	0.5	1.0
PUR250	0.44	0.14	0.5	I.0
PUR800	0.44	0.14	0.5	1.0

Table 4. Summary of analytical methods

	BNL ICP-MS	BNL FTA	LANL TIMS
Sample preparation	Calcium rhodizonate (pH > 9.5). coppt with EtOH, H_2O_2/HNO_3 microwave digestion, heat to dryness	Calcium rhodizonate (pH > 9.5), coppt with EtOH, H_2O_2/HNO_3 microwave digestion, heat to dryness	Calcium phosphate coppt from H_2O_2 solution, dissolve in 8N HNO_3
Yield monitor	^{242}Pu	Batch yield determination	^{242}Pu
Chemical separations	Fe^{2+} + NO_3^-, AG1X4, 8N HNO_3 & HCl wash, HI/HCl elution, heat to dry, 8N HNO_3	Fe^{2+} + NO_3^-, AG1X4, 8N HNO_3 & HCl wash, HI/HCl elution, heat to dry, 8N HNO,, micro anion exchange	AG1X4, 8 N HNO_3 wash, 0.36N HCl - 0.01M HF elute, electrodeposit from $NaHSO_4$ soln, strip w/HF & HNO_3, AG MP- 1, 8N HNO_3, AG MP- I, 8N HNO, wash, 0.5N HCl & HI :HCl (1:9 vol.), AG-MP- 1, H_2O_2-HCl wash, HBr elution
Source preparation	100 μl 4N HNO_3, filtered through 2 μm polypropylene filter	3 drops dried on Supmsil slides	Pu electrodeposited w/DNS from NH_4Cl soln, over deposited w/DNS
Instrumental method	HP 4500, p-concentric nebulizer, torchshield, tuned for Tl; 108.5 sec analysis	9X 1016 n/cm^2 FTA, HF etch microscope comlting	Single stage TIMS, 6 min analysis

AG1X4 = analytical grade anion exchange resin with 1X4 cross-linkage, AG MP-1 = analytical grade macro-porous anion exchange resin, coppt = co-precipitate, DNS = dihydrogen dinitro sulfato platinate (II), EtOH = ethanol, H_2O_2 = peroxide, HCl = hydrochloric acid, N = Normal, NH_4Cl = ammonium chloride, NO, $^-$ = nitrate, HNO_3 = nitric acid, N = Normal, $NaHSO_4$ = sodium hydrogen sulfate, Tl = thallium.

Dilution verification

Although dilutions of the sequential solutions could be calculated from the gravimetric determinations, it is prudent laboratory practice to verify the dilutions with measurements. It is particularly imperative to confirm the dilutions for this exercise because these are the first certified test materials of plutonium in the $nBq \cdot g^{-1}$ range. Known quantities of ^{99m}Tc were added to the plutonium reference materials, and the subsequent dilutions were confirmed by measuring the reference and diluted materials with a HPGe detector. Table 2 summarizes the results of the verification measurements for the dilutions listed in the first column. The third and fourth columns indicate the percent difference between measured and gravimetrically determined dilution factors, and the associated one sigma total propagated uncertainties, respectively. The fifth column indicates the acceptance limit for the f-test determination of any statistically significant difference between the measured and gravimetric dilution factors. As a whole, the measured dilution factors were verified to within 4%, the limit of measurement bias for the verification process. Subsequent investigations indicate the potential negative bias between the measured and gravimetric dilution factors are likely due to differences in the accuracy of Compton background subtraction for the higher activity level of the paired verification sources.

Test sample uncertainties

Table 3 summarizes the test sample uncertainties. As a result of acceptable measurement verification of the gravimetric dilution factors, the only additional uncertainty component that was propagated with the total relative uncertainty of the reference materials is from the gravimetric dilutions. These uncertainties were propagated as the root-sum-of-squares and reported as the relative expanded uncertainty for $K = 2$.[1]

Reporting format

Twenty-five samples of ^{239}Pu spiked synthetic urine of uncharacterized stability were provided to each laboratory for this study (with exception of BNL that received a set for FTA, and a second set for ICP-MS measurements). Because the long-term stability of the plutonium in the synthetic urine had not been determined, the total activity of each bottle and the massic activity concentration for each measurement result was reported. It was advised that each sample bottle be rinsed with >3M nitric acid, and the rinse solution be analyzed with the sample. The reference date and time was 12:00 noon EST, 26 February 1997.

Table 5. Relative uncertainty components for evaluation of the relative combined standard uncertainty

Uncertainty	BNL ICP-MS		BNL FTA		LANL TIMS	
statistical	($Is_{,,}$ nBq/g)	($1s_m$, %)	($1s_m$, nBq/g)	($1s_m$, %)	($1s_m$, nBq/g)	($1s_m$, %)
Concentration, nBq/g						
Blank	1.5	5900	2.5	59	3.7	80
18.5	1.4	6.4	1.1	6.0	1.0	5.5
46.3	1.2	2.6	9.3	20	3.5	1.5
148.4	0.9	0.60	43	29	24	16
271.7	1.7	0.60	61	22	78	28
Tracer	5.8%		5.8%		0.25%	
Chemical yield	9%		15%			
Other			10% (Thermal flux)		0.1% (geometry) 0.1% (spectral interference)	
Combined standard uncertainty						
Concentration, nBq/g						
Blank	5900%		62%		80%	
18.5	12%		20%		5.5%	
46.3	11%		27%		7.6%	
148.4	11%		35%		16%	
277.1	11%		29%		28%	

Table 6. Evaluation of BNL's ICP-MS results

Nuclide	NIST Values		Reported value[6] Brookhaven National Laboratory		Difference, %	ANSI N42.22 traceable	ANSI N42.22 traceability limit' %	ANSI N13.30 criteria (pass/fail)[8]	
Sample number ^{239}Pu	nBq/g[2]	Relative expanded uncertainty[3–5] ($2s_m$), %	nBq/g	Reported uncertainty ($2s_m$), %				Bias	Uncertainty
PURlOO	18.5	1.0	14.7	25	-20	Yes	±37	Pass	Pass
PUR250	46.3	1.0	40.5	22	-13	Yes	±33	Pass	Pass
PUR800	148	1.0	136	21	-8.3	Yes	±32	Pass	Pass
PUR1500	278	1.0	259	21	-6.8	Yes	±32	Pass	Pass
BLANK	0	–	0.025	12000	–	–			

Measurement techniques

NIST	Brookhaven National Laboratory
"0.1π" alpha defined-solid-angle counter with scintillation detector, 4π alpha liquid scintillation and high purity germanium counting systems	Inductively coupled plasma mass spectrometry

Table 7. Evaluation of BNL's FTA results

Nuclide	NIST Values		Reported value[6] Brookhaven National Laboratory		Difference, %	ANSI N42.22 traceable	ANSI N42.22 traceability limit' %	ANSI N13.30 criteria (pass/fail)[8]	
Sample number ^{239}Pu	nBq/g[2]	Relative expanded uncertainty[3–5] ($2s_m$), %	nBq/g	Reportecl uncertainty ($2s_m$), %				Bias	Uncertainty
PURlOO	18.5	1.0	12.7	40	-32	Yes	±60	Fail	Pass
PUR250	46.3	1.0	33.8	55	-27	Yes	±82	Fail	Pass
PUR800	148	1.0	101	70	-32	Yes	±105	Fail	Pass
PUR1500	278	1.0	212	57	-24	Yes	±86	Pass	Pass
BLANK	0		5.5	125					

Measurement techniques

MST	Brookhaven National Laboratory
"0. 1π" alpha defined-solid-angle counter with scintillation detector, 4π alpha liquid scintillation and high purity germanium counting systems	Fission track analysis

Table 8. Evaluation of LANL's TIMS results

Nuclide	NIST Values		Reported value[6] Los Alamos National Laboratory		Difference, %	ANSI N42.22 traceable	ANSI N42.22 traceability limit,' %	ANSI N13.30 criteria (pass/fail)[8]	
Sample number ^{239}Pu	nBq/g[2]	Relative expanded uncertainty[3–5] ($2s_m$), %	nBq/g	Reported uncertainty ($2s_m$), %				Bias	Uncertainty
PURlOO	18.5	1.0	34	11	82	No	±17	Fail	Pass
PUR250	46.3	1.0	74	15	59	No	±23	Fail	Pass
PUR800	148	1.0	147	31	-0.7	Yes	±47	Pass	Pass
PUR1500	278	1.0	292	57	5.3	Yes	±85	Pass	Pass
BLANK	0		46	160				–	

Measurement techniques

MST	Los Alamos National Laboratory
"0. 1π" alpha defined-solid-angle counter with scintillation detector, 4π alpha liquid scintillation and high purity germanium counting systems	Thermal ionization mass spectrometry

Reported information from the participating laboratories included:

(1) ^{239}Pu measurements in $nBq \cdot g^{-1}$ of solution; and the total combined standard uncertainty as 1s (%) (2) ±1s "standard uncertainty" components (random and systematic) which comprise the combined standard uncertainty. These may include, but are not limited to the following uncertainties: calibration factor or efficiency; dilutions or source preparation; impurity corrections; tracer calibration; gravimetric measurements; and isobaric or measurement interferences. (3) Description of how the samples were processed, including a description of solution transfer methodologies, chemical yield tracer additions, chemical separation used, measurement source preparation, and storage of samples. (4) Description of the type of measurement system used, including a general description of its operation, the type of analysis software utilized for any calculations or corrections applied to raw measurement data if applicable.

Analytical methods

Table 4 summarizes the major analytical steps used by each laboratory. The reader is advised to directly contact the participating laboratories for additional chemical separations and measurement details.

Results

Mean, standard deviation and bias

The deviations from the NIST values for each determination, the average deviation from the MST values, and the standard deviation were determined for the data that survived the outlier tests. For this study, the uncertainty was used as the indicator for the dispersion of the analysis; not as a measure of confidence in the mean value. Table 5 summarizes the participating laboratories' Combined Uncertainties ($K = 1$). The mean blank results were not subtracted from the data submitted for the spiked ranges.

Outlier tests

Because the Primary objective of this intercomparison is to evaluate the mass spectrometric and FTA technologies for their ability to measure ^{239}Pu in synthetic urine, the best reported data was used for the evaluation. Each laboratory was asked to review their data carefully for bias, and to note data of poor confidence. Those data that were noted (lost samples, below limit of detection, poor signal-to-noise ratio, or anomalous sample response to chemical separations or measurements) were reported but not used in this evaluation. The remaining data were evaluated for normal distribution. FILLIBEN's "r" criterion for goodness of fit of normal probability plots was used to detect outlying data.[2] Outlier data were also not used in this evaluation.

Report of traceability

Tables 6 to 8 summarize: each laboratory's mean bias for each certified concentration level; the determination of measurement traceability at each concentration level, per ANSI N42.22 criteria;[3] the traceability limit to which a laboratory can claim traceability (with 99.7 percent assurance); and the determination of radiobioassay measurement acceptability, per ANSI N13.30.[4]

ANSI N42.22 defines the traceability limit to MST for performance testing as:

$$|V_N - V_L| \leq 3\sqrt{(\delta_N^2 + \delta_L^2)} \qquad (1)$$

where: V_N – MST value; V_L – laboratory value; δ_N – lo total uncertainty of the MST value, V_N; δ_L – lo total uncertainty of the laboratory value, V_L

In addition, ANSI N13.30 defines criteria for acceptable bias between -25 to +50%, and acceptable uncertainty between –40 to +40%, lo total propagated uncertainty.

Analytical issues

Analytical problems

Aside from misidentified samples and computation errors, this study revealed the following analytical problems:

Analytical bias: Generally, biases approaching five percent were observed for the higher concentration test samples. It is likely that the accuracy of the chemical yield monitors (tracers) is a considerable portion of this bias. Careful preparation of yield monitors should remove most of the analytical bias. In addition, FTA was handicapped with bias limitations when track density was high, and when batch chemical yield corrections were used.

Uncertainties: BNL's FTA and LANL's TIMS relative uncertainties increased with increasing plutonium concentration. This is contrary to intuition and should be investigated for root cause by each laboratory.

Imprecision: Most of the poor uncertainty is caused by high variable blanks and low chemical yield (see below). Large measurement uncertainty could result in failing the ANSI N13.30 criteria for precision.

High variable blank: LANL results suffered from high and variable blanks. LANL's high and variable blanks are probably the result of contamination of the samples during initial sample dissolution and chemical separation in a general laboratory prior to the final chemical separations, mounting and TIMS

measurements in their clean-laboratory facility. BNL's ICP-MS and FTA results, on the other hand, had very low and consistent blanks. Presumably, BNL has developed extraordinary cleanroom techniques and used ultra-pure reagents to minimize sample and reagent contamination. The results of this study indicate that BNL's performance may be linked to a laboratory's ability to control and minimize any blank contributions. LANL should undertake careful study of their analytical system to identify and control sources of interferences at these test levels of ^{239}Pu activity.

Low chemical recovery: Discussions with the investigators indicated that the chemical recovery of Pu from natural mine samples are typically in the 70–80% range. The synthetic urine used in this study caused chemical recoveries to occasionally decrease to 20%. The root cause for such low chemical recoveries need to be investigated, particularly because it causes this technology evaluation to be inaccurate (particularly the evaluation of uncertainty and MDA). No decrease in chemical recovery has been previously reported for alpha-spectrometric analyses of synthetic urine (same recipe) used as part of the radiobioassay DOELAP.

Lost data: 21% of the reported results were not included in the study because they were reported as less than MDA, were analytical outliers, poor uncertainty, biased by excessive overlapping tracks, or of poor reliability. Reliability of the analytical systems must be characterized through systematic methods evaluations at each participating laboratory and statistically evaluated.

Study limitation

A serious limitation to this study is the absence of an evaluation of important isobaric and chemical interferences in the synthetic urine matrix. Addition of interference analysis would have also meant an examination of chemical separations and measurement selectivity. Interferences that may be present in natural mine include calcium, iron, lead, uranium and thorium isotopes, ^{240}Pu and ^{241}Pu. The results of this study should be interpreted as being collected under optimum conditions. Including interference data would have more closely simulated true analytical performance on natural urine.

An additional limitation of the study was instruction to analyze the whole sample. BNL's FTA was optimized for ultra-low levels of ^{239}Pu measurements and the higher concentration samples resulted in samples with very dense fission tracks. The subsequent track overlap generally causes a negative bias in the results. Although the highest concentration samples were identified to BNL for FTA measurements to allow appropriate accommodation for this source of bias, it would have also been advisable that BNL be allowed to have taken smaller subsamples.

Despite these study shortcomings, sufficient data exists to address the underlying objectives of this study, as provided in the next section.

Discussion

Uncertainty

Expanded combined uncertainties ($K = 3$) are listed in Tables 6 to 8. BNL ICP-MS had the best uncertainty among all of the measurements. LANL TIMS had wider uncertainty (by factors of about 1.5 and 4, respectively). It is likely that the analytical blank control by BNL played a key role in their well controlled performance for measurement uncertainty. Although LANL ran internal blank controls that were used to correct the submitted values, the results for the unspiked synthetic urine samples indicated additional sources of instrument signal.

BNL FTA's uncertainty was 2-3 times larger than for its ICP-MS. This is because there are inherent uncertainty limitations for FTA when there are: (a) few tracks, resolution is good but there is poor uncertainty, (b) many tracks, resolution is poor from track overlap, and uncertainty and bias are adversely affected, and (c) use of a moving batch recovery correction factor technique to correct for chemical yields.

It is unclear at this time why the LANL TIMS relative measurement uncertainty increased as the plutonium concentration increased. In general, the reverse is expected because of higher ion fluxes. This point is left for future investigations, including the effects of incomplete chemical separations and isobaric interferences.

Bias

Tables 6 through 8 lists the percent bias from the NIST values at each ^{239}Pu concentration level. Interpretation of these results is complicated by measurements having poor uncertainty. However, it is clear that BNL ICP-MS had the smallest bias values. BNL ICP-MS unambiguously demonstrated measurement capabilities for ^{239}Pu at the μBq level with high measurement bias and uncertainty. The agreement with the NIST values lends support to the assumption that the test samples were stable and accessible during this study. The BNL ICP-MS value for the blank samples was extremely low, and was probably responsible for the good performance. It is noted, however, that there is a systematic negative bias. It is left to future investigations to determine if the negative bias was due to a systematic difference in the certification of the ^{242}Pu tracer.

Although LANL TIMS had larger bias at the 18.5 and 46.3 nBq·g^{-1} levels, excellent bias values were obtained for their routine bioassay program at the 148 and 278 nBq·g^{-1} levels, although with poorer uncertainty. These results also illustrate that TIMS capabilities could be improved for measurements down to 20 nBq/g samples when the unaccounted blank contributions are controlled.

The BNL FTA bias is larger than those from ICP-MS, but are somewhat smaller than LANL TIMS at the lower concentrations and poorer at higher concentrations. As mentioned before, the poorer FTA performance is related to track density and use of a moving batch recovery factor technique to correct for chemical yield. It would be possible for BNL FTA to improve its bias performance when internal tracers (chemical yield monitors) are used. The results of this study indicate that FTA can make measurements within about 80% of a true value, 99.7% of the time, over the ^{239}Pu 3.7-55.6 μBq range.

ANSI performance criteria

The results from the laboratories were evaluated against ANSI N42.22 and ANSI N13.30 performance criteria. ANSI N42.22 defined acceptable verification of measurements (measurement traceability testing) as "the difference between the NIST value and the (test) value of the manufacturer be less than the total propagated uncertainty of the difference with a coverage factor of three." ANSI N13.30 defined acceptable relative bias (-25 to +50%) and relative uncertainty (±40%, $K=1$) criteria for radiobioassay measurements.

All laboratories demonstrated their ability to make traceable measurements, per ANSI N42.22 criteria, at the 148.4 and 277.7 n·Bq·g^{-1} concentration levels. Additionally, at the 18.5 and 46.3 n·Bq·g^{-1} concentration level, BNL's ICP-MS and FTA measurements were traceable. FTA's ability to make traceable measurements at the lower concentration levels, however, was primarily due to relatively large total propagated uncertainties.

All laboratories passed both the uncertainty and bias ANSI N13.30 criteria at the 277.7 n·Bq·g^{-1} level. Furthermore, BNL's ICP-MS satisfied the ANSI N13.30 criteria at all four concentration levels.

Minimum detectable amount (MDA)

The estimated MDAs were derived from the ANSI N13.30 equations; the simplified equation, $MDL=4.65s_b+3$, was not used because of significant contributions from systematic biases.

The general MDA equation from ANSI N13.30, when α and β are equal, is:

$$MDA = \frac{(1+\Delta_K)(2\Delta_B B + 2ks_0 + 3)}{KT} \qquad (2)$$

where B – the total count of the appropriate blank, s_0 – the standard deviation in the net sample count with no additional analyte, defined by ANSI N13.30 Eq. (2), K – calibration factor, (including correction for self absorption when appropriate), Δ_K – the maximum fractional systematic error bound in the calibration factor K, (like Δ_B, Δ_K cannot be estimated using replicate measurements, and must be estimated by the professional judgment of the analyst), Δ_B – the maximum expected fractional systematic error bound in the appropriate blank, (the factor of 2 before Δ_B takes into account the maximum systematic error bound when the background and sample measurement errors are of opposite sign), k – the abscissa of the standardized normal distribution corresponding to the 0.05 probability level, for $\beta = 0.05$ and $\alpha = 0.05$, $k = 1.645$, T – standard counting time for the procedure.

The MDA can be obtained from data in units of count-rate from:

$$MDA = (1+\Delta'_k)(2\Delta_B B' + 2ks'_0 + 3)/K' \qquad (3)$$

where $B' = BIT$, $s'_0 = s_0/T$, $K' = K/T$, $\Delta'_k = \Delta_k$ since they both represent the same fractional systematic relative fixed error.

The unprimed quantities are used when total counts are used in the computation, and the primed quantities are used when the count rates are computed.

For this intercomparison Equation seven, of ANSI N13.30, was used to calculate MDAs. It was further assumed that $\Delta'_k = \Delta_B$. The BNL ICP-MS MDA at "0" analyte concentration was calculated as a standard deviation because the estimated MDAs were independent of sample concentration. On the other hand, because several blank sample results were not reported, and estimated uncertainties for the blank sample results were large and varied with sample concentration, an extrapolation method was chosen to improve the reliability of estimating the BNL FTA and LANL TIMS ICP-MS MDA's at "0" analyte concentration. MDA's were calculated at each concentration level, and extrapolated back to "0" n·Bq·g^{-1}. The reported MDA's for this study are the mean values, and the confidence intervals (CI's) are reported as two standard deviations of the calculated MDA's.

In general, routine alpha-spectroscopy's MDA is about $3 \cdot 10^6$ n·Bq. These results indicate that mass spectroscopy's (BNL ICP-MS and LANL TIMS) MDA is about 3000 times lower than for alpha spectroscopy. This is in contrast to the study by LEE et al.[5] and HUTCHINSON et al.[6] where they found ICP-MS *only* had comparable measurement capabilities to alpha-spectroscopy for ^{239}Np in synthetic urine. FTA's MDA is comparable to that expected from ICP and thermal ionization mass spectroscopy, but is a factor of 10-100 times less certain.

Table 9. Estimated MDA's for 200 g samples

Laboratory	MDA, nBq/200 g sample	*95%* Confidence interval, %
BNL ICP-MS	***1600***	***35***
BNL FTA	1200	***1900***
LANLTIMS	600	***590***

The best MDA's were obtained by BNL's ICP-MS and LANL's TIMS. Although the estimated BNL MDA is somewhat larger than LANL's, it is known with much better uncertainty and confidence. The estimated BNL FTA MDA was determined with poorer uncertainty.

Technical issues

The results of this study can now be used to address the technical issues raised during the design of the study protocol.

Stability of the plutonium in glass bottles and in the synthetic *urine:* As a minimum, over the short-term of a few weeks and by washing the bottle with strong acid, the plutonium appears to be stable in the glass bottles and in the synthetic urine. If all the bias were due to a Pu stability issue, then the BNL ICP-MS results indicate stability of the test samples to better than 8% at the 148-278 $n \cdot Bq \cdot g^{-1}$ levels, and better than 20% at the 15–41 $n \cdot Bq \cdot g^{-1}$ levels.

Contamination from plutonium in the reagents used to *make the synthetic urine:* The BNL ICP-MS and FTA results indicate contamination of the test samples by plutonium in chemical reagents to be negligible ($<6\ n \cdot Bq \cdot g^{-1}$, and probably as low as $\approx 0.03\ n \cdot Bq \cdot g^{-1}$).

Adequacy of the synthetic urine as a substitute for natural urine: The ANSI N13.30 standard allows use of synthetic urine as a test matrix, synthetic urine was used for pilot testing the efficacy of the ANSI N13.30 standard, and synthetic urine will be used for the radiobioassay DOE Laboratory Accreditation Program. However, it was pointed out by all of the participating laboratories that chemical yields were substantially lower than anticipated. For example, LANL reported chemical yields as low as 20% – their average chemical yield for radiourine assay is 80%. The low chemical yield substantially lowered analytical sensitivity and increased measurement uncertainty. A systematic study will be necessary at each laboratory to optimize chemical yield from synthetic urine analysis. It is likely, however, that the resulting analytical protocol may be substantially more robust from that in daily use for natural urine. However, the results of this study provides a baseline lower limit to mass spectrometry's capabilities, against which improvements can be measured.

Summary and outlook

The results of this study revealed current ICP-MS, TIMS, and FTA capabilities. These are summarized in Table 9. It is apparent that mass spectrometry is currently capable of successfully competing with FTA's sensitivity, and at considerably higher uncertainty down to the 3.7 μBq/sample level. In addition, mass spectrometry has the potential to improve with new technology (e.g., new nebulizer designs, multi-pulse detection systems, selective laser ionization) to provide more accurate and precise measurements than FTA. Chemical technologies can be improved with robotics, use of new column chromatography techniques, and the savings in terms of human resources can be shared by all three measurement technologies. When attention is given to careful chemical separation and purification, accurate and precise mass spectrometric measurements of μBq quantities of ^{239}Pu in synthetic urine are achievable. Furthermore, it is feasible that mass spectrometry can provide higher sensitivity measurements of ^{239}Pu for routine occupational radiobioassay programs and emergency assessment evaluations than alpha-spectrometry.

In summary:

ICP-MS results indicated the tremendous ability to accurately and precisely measure μBq quantities of ^{239}Pu in synthetic urine, while providing the same MDA level as FTA.

FTA can also measure μBq quantities of ^{239}Pu in synthetic urine, but with considerably larger uncertainty than mass spectroscopy.

TIMS can be used for routine radiobioassay programs and has the potential to provide better measurement capabilities than FTA, but a considerable effort must be made to identify and control root causes of high blanks and imprecision.

Controlling analytical blank is crucial for measuring ultra-low levels of ^{239}Pu in urine, which also means careful and exhaustive chemical separations cannot be avoided.

The chemists must find ways to improve chemical yields to further enhance measurement sensitivity and reliability.

Conclusions

The prime objective of this study was to assess the current capabilities of FTA, ICP-MS and TIMS to measure μBq quantities of ^{239}Pu in urine. It is clear that all three methods have the capabilities to make such measurements. Programs such as the US DOE's effort to

carefully monitor the plutonium dose to the Marshallese people can be assured that the potential to measure μBq levels of ^{239}Pu with sufficient bias and uncertainty has been demonstrated. BNL's ICP-MS work demonstrated that accurate and precise measurements are already a reality. This reality, however, is probably dependent on the laboratory's ability to minimize and control the analytical blank. Such control may be achieved with highly skilled professionals, in dedicated ultra-clean laboratory facilities, with ultra-pure reagents. Measurements of such small quantities of plutonium is technically difficult, and lost data (21% in this study) or repeat analysis must be minimized with robust analytical and measurement procedures. Although FTA does not have the analytical bias of high quality ICP-MS, this study has demonstrated that it potentially has comparable MDA to ICP-MS. Unless the inherent disadvantages of FTA (batch yield correction, track overlap, and poor statistics) can be overcome, it maybe advantageous that a larger share of development resources be focused on mass spectromeaic analyses. TIMS was found to be able to provide excellent results above 30 μBq ^{239}Pu/sample for routine radiobioassay programs, and could be improved with better blank control. LANL TIMS could enhance their capabilities considerably through minimization and control over analytical blank, higher chemical recovery, improved uncertainty, and smaller bias yield monitors. With future improvements in technology and techniques, it is anticipated that ICP-MS and TIMS will satisfactorily meet the ANSI N42.22 criteria for traceability and the ANSI N13.30 criteria for bias and uncertainty, even at μBq levels of plutonium in the complex urine matrix.

Secondarily, the technical issues of test sample preparation and stability have been addressed. This study has demonstrated that careful serial dilutions of the plutonium SRM over nine orders of magnitude to $nBq \cdot g^{-1}$ concentrations can be done accurately, that the dilutions can be confirmed by measurement to within a few percent, and the plutonium in synthetic urine remains stable and accessible for analysis (to within 5%) for at least a few weeks. The results of this study confirms the efficacy of the protocol to prepare these test materials.

Recommendations

It is recommended that improvements to the bias, uncertainty and sensitivity of plutonium in urine mass spectrometry metrology could assist in occupational worker's health and safety, validation of excretion models; identification of the source-terms; and litigation dispute resolution, be initiated by:

(1) Evaluating and contrasting techniques to determine the critical elements controlling blank contamination and variability, chemical yield, measurement imprecision; and analytical bias.

(2) Developing a consistent method to calculate FTA and ICP-MS measurement uncertainties and detection limits.

(3) Preparing standard reference materials of ^{239}Pu, and ^{242}Pu chemical yield tracer, at 11.1 $\mu Bq \cdot g^{-1}$ level in high purity 1N HNO_3 (anion exchange separated) for use by the mass spectroscopy community.

(4) Conducting intercomparison of ^{239}Pu in the range of 3-56 μBq (100-1500 aCi) per 200 ml of synthetic urine containing chemical and isobaric interferences: ^{240}Pu, ^{241}Pu uranium and thorium isotopes, and trace-elements to more carefully test ICP-MS, TIMS and FTA under more realistic conditions.

References

1. B. N. TAYLOR, C. E. KUYATT, Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results, NIST Technical Note 1297, National Institute of Standards and Technology, Gaithersburg, MD, 1994.
2. J. J. FILLIBEN, Technometrics, 17 (1975) 111.
3. American National Standards Institute, Traceability of Radioactive Sources to the National Institute of Standards and Technology (NIST) and Associated Instrument Quality Control, ANSI N42.22-1995.
4. American National Standards Institute, Performance Criteria for Radiobioassay, ANSI N13.30-1996.
5. S. C. LEE, J. M. R. HUTCHINSON, K. G. W. INN, M. THEIN, Health Phys., 68 (1995) 350.
6. J. M. R. HUTCHINSON, S. C. LEE, K. G. W. INN, An Jntercompruison Study of Neptunium-237 Determination in Artificial Urine Samples, Report to DOE DE-AIO5-91OP21969, 1993.